

A Natural Experiment on Discrimination in Elections*

David E. Broockman
UC Berkeley

Evan J. Soltas
MIT

May 5, 2020

Abstract

We exploit a natural experiment to study discrimination in elections. In Illinois Republican presidential primaries, voters vote for delegates bound to presidential candidates, but delegates' names convey information about their race and gender. We identify discrimination from variation in vote totals among delegates bound to the same presidential candidate and who face the same voters. Examining delegate vote totals from 2000 to 2016, we estimate nonwhite delegates receive 9 percent fewer votes. We find essentially no gender discrimination. Negligible incentives for statistical discrimination, costs to preferred presidential candidates, and heterogeneity are consistent with an interpretation of this behavior as taste-based.

Keywords: Racial Discrimination, Voter Behavior, Taste-Based Discrimination

JEL Codes: D72, J15

*Broockman: UC Berkeley (dbroockman@berkeley.edu). Soltas: MIT Department of Economics (esoltas@mit.edu). We are grateful to Dave Wasserman for the news article that served as inspiration for this study. We thank the editor Thomas Fujiwara and the anonymous referees for their comments and suggestions. For helpful feedback, we thank Avi Acharya, Hunt Allcott, David Autor, Tim Besley, Kate Casey, Stefano DellaVigna, Dana Foarta, Matthew Gentzkow, Don Green, Dan Ho, Saumitra Jha, Ilyana Kuziemko, Horacio Larreguy, Michael Luca, Jon Levin, Pablo Montagnes, Sergio Montero, Nicola Persico, Ken Shotts, Erik Snowberg, Jorg Spenkuch, Seth Stephens-Davidowitz, Francesco Trebbi, Ebonya Washington, Ismail White, and seminar participants at LSE, the NBER, Oxford, POLMETH, Queen's University, SPSA, Stanford University, UC Berkeley, UC Davis, UC Riverside, the University of Toronto, the University of Zurich, the University of Southern California, the Wallis Institute, and Yale. For their able research assistance, we thank Hadil Almowafak, Julia Espero, Corey Garcia, Chris McConnell, Leela Srinivasan, KS Tey, and Yao Yang. Soltas gratefully acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. 1122374.

1 Introduction

Racial and ethnic minorities and women are underrepresented among elected officials in many countries.¹ A large body of research indicates that such underrepresentation contributes to disparities by race, ethnicity, and gender in a variety of economic and political outcomes. Electing minority or female officials has been found to reduce gaps in public goods provision (e.g., [Pande, 2003](#); [Chattopadhyay and Duflo, 2004](#)), resulting in relative gains in health, education, and criminal justice ([Beaman et al., 2012](#); [Clots-Figueras, 2012](#); [Iyer et al., 2012](#); [Fujiwara, 2015](#)). One potential explanation for the underrepresentation of minorities and women in elected office is voter discrimination, wherein voters are less likely to vote for a minority or female candidate than an otherwise-identical non-minority or male candidate.

Understanding whether voters discriminate by race and gender and the mechanisms for any such discrimination are central questions in political economy and would inform significant policy debates. Research finds that individuals engage in racial discrimination in product and labor markets (for review, see [Bertrand and Duflo, 2017](#)), but credible evidence on whether they do so when voting is limited. On average, minorities run in different electoral districts, at different times, on different platforms, with different party affiliations, for different offices, with different pre-election experience and campaign resources, and so on, making credible identification of voter discrimination difficult in most electoral settings ([Stephens-Davidowitz, 2014](#)). In settings beyond politics, researchers have identified discrimination with a variety of credible research designs.² However, it is challenging to adapt many of these strategies to elections.

In this paper we analyze a natural experiment to study voter discrimination against nonwhite and female political candidates. This natural experiment occurred in four recent Illinois Republican presidential primary elections: 2000, 2008, 2012, and 2016.³ Unique features of the institutional environment means discrimination in these elections can plausibly be causally identified. For voters to fully express their preferences in Illinois Republican presidential primary elections, they must vote for multiple individual candidates for delegate to the Republican National Convention who appear on the primary ballot. In each congressional district in Illinois, there is a fixed number $N \in \{2, 3, 4\}$ of delegate candidates

¹For example, in the United States as of 2015, 38 (51) percent of the population was nonwhite (female), compared to 17 (20) percent of the U.S. Congress ([Manning, 2016](#)). For brevity, in the paper we refer to non-Hispanic whites only as whites, and all other groups as nonwhites. We often use “racial discrimination” as a shorthand for racial and ethnic discrimination.

²These include audit studies (e.g., [Pager et al., 2009](#)), correspondence studies ([Bertrand and Mullainathan, 2004](#)), natural experiments ([Goldin and Rouse, 2000](#); [Price and Wolfers, 2010](#)), and lab and field experiments (for review, see [Guryan and Charles, 2013](#); [Bertrand and Duflo, 2017](#)).

³We use these years because relevant data was not available prior to 2000 and the 2004 Republican presidential primary was not contested, as President George W. Bush was running for renomination unopposed.

who appear on the ballot for each presidential candidate, and voters cast votes for up to N delegates. The top N vote-getting delegates in each district win and are bound to vote for their presidential candidate at the convention. However, the delegates’ names also appear on ballots, and the delegates’ names convey information about their race and gender.

To identify racial and gender discrimination in this setting, we exploit the fact that we observe the vote totals of multiple delegates with the same platform—that is, bound to the same presidential candidate—on the same ballots, and voted on by the same voters, all of whom voters should select to fully support their preferred presidential candidate. Our identification strategy is to examine variation in delegate vote totals by delegate race and gender within such groups of delegates. For example, suppose the delegates for Mitt Romney in Illinois’ first congressional district in 2012 were named Tom, Dick, and José. To maximize the value of their ballot, a Romney supporter should cast their three votes for Tom, Dick, and José. However, to the extent Romney supporters engage in racial or ethnic discrimination, some may vote for Tom and Dick but not for José, leaving José’s vote totals lower than Tom and Dick’s vote totals. We observe 816 unique natural experiments of this form.

This election design has important advantages for studying discrimination. In typical elections, voters may value myriad dimensions of candidates—such as ideology, competence, or past performance—many of which may correlate with candidate demographics and few of which are perfectly observable. Here the voter’s problem is dramatically more straightforward: For voters seeking to fully support their preferred presidential candidate, a delegate candidate’s sole relevant dimension is the presidential candidate to whom they are bound,⁴ which is clearly printed on ballots, such that both voters and researchers can perfectly observe it. White and male delegate candidates running alongside nonwhite and female delegate candidates on the same platform, for the same office, on the same ballots, in front of the same voters therefore provide a naturally-occurring control group that allow us to rule out factors that would confound other research designs.

Analyzing variation in delegate vote totals among delegates bound to the same presidential candidate and who appear on the same ballots in front of the same voters, we find that delegates receive approximately 9 percent fewer votes when their names indicate they are not white. We also find, however, essentially no discrimination against women: Delegates whose names indicate they are female receive on average about the same number of votes as delegates whose names indicate they are male, if not slightly more in some specifications.

⁴In Section 4.3, we consider other dimensions voters may value besides the presidential candidate to whom delegates are bound, such as if a delegate is an existing elected official, is a “local notable,” or is listed higher on the ballot. We find these dimensions are uncorrelated with delegate race. Serving as a delegate most resembles hobbyist consumption, rather than a career investment. See Appendix H for discussion, including quotes from convention delegates about why they volunteered to attend.

Examining the results by ethnicity, we find clear evidence of discrimination against delegates whose names indicate they are Hispanic, East Asian, Middle Eastern, or Indian. Our estimates for discrimination against black delegates are similar, although there are few clearly black delegates, making our estimates of anti-black discrimination less precise.

We next consider what theoretical mechanisms could account for this discrimination. Discrimination may be taste-based—wherein voters act as if they have preferences over candidate race and gender (Becker, 1957). It may also be statistical—wherein voters accurately use candidates’ race and gender to infer non-racial or non-gender dimensions such as ideology and can advance these non-demographic preferences by discriminating (Phelps, 1972; Akerlof, 1976). Taste-based discrimination is of particular interest because it implies that voters act as if they pay a “psychic cost” (Becker, 1957) of voting for candidates from disfavored demographic groups and accept trade-offs on candidates’ non-demographic dimensions to avoid paying these “psychic costs.” As compared to markets (List, 2004), taste-based discrimination may be especially likely in elections because individual voting decisions are usually inconsequential for outcomes. Nevertheless, taste-based discrimination may be consequential in aggregate if many voters engage in it, reducing both minority representation and the appeal of election winners on other dimensions.

Institutional features of the setting we study suggest tastes—and, in particular, the “psychic costs” of voting for delegates of a disfavored race or gender—as the likely mechanism for the discrimination we detect. In most electoral settings it would be difficult to distinguish between taste-based and statistical discrimination.⁵ Taste-based discrimination occurs when individuals behave as if they prefer candidates inferior on non-racial or non-gender dimensions in order to avoid incurring “psychic costs” from selecting otherwise-preferred candidates with disfavored demographics. Voters must behave in precisely this manner to discriminate against delegates in this environment. If voters do not vote for all their preferred presidential candidates’ nonwhite delegates, this advantages delegates for presidential candidates they prefer less, undermining the nomination prospects of voters’ preferred presidential candidates. In addition, although we cannot rule out all alternative interpretations, incentives for statistical discrimination should be naturally absent. Under convention rules, delegates have essentially no discretion. Even if a rational voter were unaware that delegates had no discretion, to engage in statistical discrimination, she would need to maintain beliefs we view as implausible: that nonwhite delegates bound to her presidential candidate of choice would be less likely to vote for the voter’s presidential candidate of choice at the

⁵Existing approaches in the gender literature include comparisons of vote totals controlling for observables, survey-based experiments (e.g., Teele et al., 2018), and testing implications of voter bias on politician quality (Anzia and Berry, 2011; Ferreira and Gyourko, 2014; Vogl, 2014).

convention than white delegates bound to an opposing presidential candidate.⁶

Heterogeneity in the magnitude of discrimination across candidates, elections, and geography is also consistent with predictions of a taste-based interpretation. Most significantly, consistent with taste-based discrimination also affecting voters’ choices of presidential candidates, and not only delegate candidates, we estimate that voters for nonwhite presidential candidates harbor significantly weaker racially-discriminatory tastes than voters for white presidential candidates. We also find that voters for a female presidential candidate actually discriminate in favor of female delegate candidates. These results are consistent with voters’ racial and gender tastes having stakes for their choices of presidential candidates, as voters appear to endogenously select out of voting for nonwhite and female presidential candidates in a manner strongly correlated with our estimates of their collective tastes. In addition, consistent with [Becker \(1957\)](#), discrimination also appears to decrease when it is more costly to voters’ preferred presidential candidates: Voter discrimination against nonwhites is less when voters are more likely to be decisive, although we find it still appears to persist even in the most competitive elections. The geographic areas where discrimination appears strongest is also in line with expectations from prior research.

We present a variety of robustness checks on our results. We show the results are similar when we use each of three different strategies to measure the racial signals delegates’ names send voters: a measure based on the background of others with their last name in data provided by the U.S. Census, a measure based on anthropological data about the etymology of their full names, and a measure based on Americans’ subjective perceptions based on their full names. We further show that ballot-order effects do not drive the results. In addition, we construct three measures of possible prior information voters could have about delegates; our results are robust to excluding delegates about whom voters might have had other information and to controlling for this information. We also consider alternative interpretations of the results, such as residual incentives for statistical discrimination, voter misunderstandings about the primary, voter signaling to presidential candidates or party elites, voter indifference across presidential candidates, and voter inferences about presidential candidates on the basis of nonwhite delegates. Although we cannot rule out alternative mechanisms beyond taste-based discrimination related to voter misunderstanding and signaling, [Section 4.4](#) and

⁶It is therefore unlikely that voters with mistaken beliefs about the primary and delegates would perceive incentives for statistical discrimination. We consider the plausibility of statistical discrimination in more detail in [Sections 2](#) and [4.4](#), as well as in [Appendix J](#). Under convention rules, even if delegates fail to appear or fail to cast their vote for the presidential candidate to whom they are bound, their vote is counted for the presidential candidate to whom they are bound regardless. Moreover, [Becker \(1957\)](#) defined taste-based discrimination to encompass inaccurate beliefs about minorities: “An employer may refuse to hire a [black person] solely because he erroneously underestimates their economic efficiency. His behavior is discriminatory...[A] taste for discrimination incorporates both prejudice and ignorance” (p. 16–17).

Appendix J provide further evidence on and discussion of these alternatives.⁷

We find voter racial discrimination likely had aggregate consequences for delegate election outcomes. It is likely that discrimination against several presidential candidates’ nonwhite delegates reduced their vote totals sufficiently that white delegates for less-preferred presidential candidates won and served instead.⁸ This illustrates the two consequences of taste-based racial discrimination in elections appear to have been realized in this environment: In reducing nonwhite representation, voters who discriminated also elected candidates less appealing to them on other dimensions. To contextualize the magnitude of our point estimate, we also offer back-of-the-envelope calculations that apply our estimate to Republican primary elections for the U.S. House of Representatives. Discrimination of this magnitude would decrease the share of nonwhites in the U.S. House by about 3 percentage points. Voter discrimination against racial-minority candidates therefore plausibly contributes to their underrepresentation in government, which other research has found lies at the root of important social disparities.

Like much previous research on discrimination (e.g., Goldin and Rouse, 2000; Price and Wolfers, 2010; Doleac and Stein, 2013; Rubinstein and Brenner, 2014; Glover et al., 2017), we exploit an institution whose unique properties facilitate otherwise-elusive causal inference. Therefore, we carefully consider the generalizability of our findings as to how detecting discrimination in this setting should inform views of the plausible extent of discrimination in other elections. The underrepresentation of minorities and women in the U.S. is most extensive among Republican elected officials, making Republican primaries—the de-facto elections in about half of U.S. electoral districts—of particular substantive importance. Our finding that voters appear to endogenously select into voting for presidential candidates in a manner strongly correlated with our estimates of their collective tastes is consistent with voters’ racial and gender tastes having stakes for their choices of presidential candidates. However, discrimination could be greater in other primaries where voters may have less information or weaker preferences. The presidential primaries we study are relatively high-stakes elections, determining the Republican presidential nominee, and where evidence indicates voters have strong candidate preferences. In addition, our estimates can capture only the “psychic cost” of voting for nonwhites and women, not any “psychic costs” of being represented by them nor statistical discrimination against them. Accounting for such distinctions, the total disadvantage for nonwhite candidates in other elections due to voter discrimination may well

⁷There we discuss in more detail why voter misunderstanding should not produce incentives for statistical discrimination. We also fielded a survey of Illinois Republican primary voters that found limited voter misunderstanding or perceived incentives for statistical discrimination. However, 10 percent said they felt “uncomfortable” voting for nonwhites, consistent with taste-based discrimination.

⁸In Section 3.6 we discuss why presidential campaigns may nominate nonwhite delegates despite this cost.

be larger. On the other hand, discrimination could be smaller in general elections where partisan preferences may be more important, or in elections where voters have on average weaker racial tastes.

The difference in our results for race versus for gender merits an additional comment. In particular, our finding that voters do not appear to discriminate against women is consonant with evidence from gender quotas (Baltrunaite et al., 2014, 2019; Esteve-Volart and Bagues, 2012; Casas-Arce and Saiz, 2015; Besley et al., 2017) and survey experiments (Schwarz and Coppock, 2019) that suggest political institutions, rather than voters, as the principal obstacle to increasing female political representation in Western democracies. On the other hand, the discrimination we detect against nonwhites fits with Washington (2006), who shows that a fraction of white voters turns out specifically to vote against nonwhite U.S. House candidates. Our results therefore lend support to claims in the literature on gender that underrepresentation likely results from different combinations of contributing factors for nonwhites and women.

We explain the context and natural experiment in greater detail in Section 2. Section 3 introduces our main data sources. Section 4 presents our empirical strategy, results, and robustness checks. Section 5 examines the pattern of heterogeneity in discrimination and argues it is consistent with a taste-based interpretation. Section 6 discusses implications for other elections, U.S. election law, and strategies for increasing nonwhite and female representation.

2 The Illinois Republican Presidential Primary

2.1 Why Study Republican Primaries?

Why are nonwhites and women underrepresented among U.S. officeholders? Stylized facts about U.S. politics suggest discrimination among Republican voters in primary elections may play a significant role. First, U.S. voters have strong partisan preferences, with relatively weaker preferences among candidates of the same party (Green et al., 2002), implying that any taste for a candidate’s race, gender, or ethnicity may be especially determinative in primary rather than general elections. Second, white Republican voters have been more racially conservative than white Democratic voters since the 1960s civil rights realignment (Kuziemko and Washington, 2018). Nevertheless, Republican primaries constitute the de-facto election in about half of U.S. electoral districts, where Republican nominees reliably win in general elections. To the extent voter taste-based discrimination affects the demographic composition of U.S. officeholders, one thus might expect it to do so especially through Republican

Table 1: Racial and Gender Composition of Officeholders and Voters by Party

	% Nonwhite		% Female	
	Republicans	Democrats	Republicans	Democrats
U.S. House Members	4%	31%	9%	31%
Primary Election Voters	11%	35%	51%	60%
Party Identifiers	15%	43%	53%	59%
General Election Voters	23%		54%	

Notes: This table reports the national shares nonwhite and female of U.S. House members, primary voters, party identifiers, and general election voters who are Republicans or Democrats. Data on the racial and ethnic composition of U.S. House Members from 2006-2014 was collected by [Fraga \(2013\)](#). Data on the gender composition of U.S. House Members was collected by the [Center for American Women and Politics \(2016\)](#). Data on the racial, ethnic, and gender composition of primary-election voters, general-election voters, and party identifiers is from the 2016 Cooperative Congressional Election Study (CCES), a large survey of American voters ($N = 64,600$) ([Ansolabehere and Schaffner, 2017](#)). Whether CCES participants voted is measured from administrative records. Party identification is measured by the survey question “Generally speaking, do you think of yourself as a...?” with answers that include “Democrat” and “Republican.”

primary elections.

Consistent with this possibility, Table 1 shows that both nonwhites and women are underrepresented among Republican U.S. House members relative to the populations of Republican primary voters and adults who identify as Republicans. These disparities are present but notably smaller among Democrats. Discrimination in Republican primaries could contribute to these descriptive patterns because, in a considerable share of U.S. electoral districts, whoever the Republican Party nominates is likely to win the general election. Any racial, ethnic, or gender discrimination in Republican primaries in such Republican-leaning districts may therefore alter the demographic composition of U.S. elected officials. While of course other plausible explanations exist for these descriptive patterns, the natural experiment we analyze provides new evidence consistent with voter racial discrimination in Republican primaries as a contributor.

2.2 Design of the Primary

We study Illinois Republican primary elections, taking advantage of their unique design. The “delegate loophole primary” design of the Illinois Republican presidential primary is unique within the United States.⁹ Voters vote separately for some number of delegates—

⁹Delegate loophole primaries were once common in the U.S. but were largely replaced by candidate-based primaries in the 1970s as part of reforms intended to empower voters in primary elections ([Shafer, 1988](#)). Difficulty in locating election returns in other states prior to the 1970s, and the low number of nonwhite delegate candidates who are likely to have run at that time, limit our ability to extend our analysis to these earlier primaries. Some other states also list delegates on ballots, but in none of these other states does

Figure 1: Section of Sample Ballot

**FOR DELEGATE TO THE
NATIONAL NOMINATING CONVENTION
EIGHTEENTH CONGRESSIONAL DISTRICT**

(PLEASE NOTE: Next to the name of each candidate for
delegate appears in parentheses the candidate's preference for
President of the United States or the word "uncommitted".)
(Vote for not more than three)

- ☐ **ROBERT BROWNING** (CHRISTIE)
- ☐ **MARY K. BROOKHART** (CHRISTIE)
- ☐ **DONNA K. THOMPSON** (CHRISTIE)
- ☐ **JIM EDGAR** (BUSH)
- ☐ **BILL BRADY** (BUSH)
- ☐ **RAYMOND POE** (BUSH)
- ☐ **KENT GRAY** (TRUMP)
- ☐ **SANDRA YEH** (TRUMP)
- ☐ **WILLIAM GRAFF** (TRUMP)

Notes: This figure shows a relevant delegate-selection section of the 2016 Republican primary ballot from McLean County in Illinois' 18th Congressional District. See Appendix Figure A1 for a copy of the full ballot.

usually three—who are bound to represent a given presidential candidate at the Republican National Convention. If elected, the delegate candidates have essentially no discretion in the votes they cast at the Republican National Convention. Rule 16(a)(2) of the 2012 and 2016 Republican National Conventions specifies that delegates who do not vote for the candidate to which they are bound have that vote canceled, and the Secretary of the Convention records the vote as for the candidate to which the delegate was bound. Additional features of these elections, reviewed in Section 4.4, make it unlikely that even voters unaware of this rule would perceive incentives for statistical discrimination. In all the elections we study, about 80 percent of the Illinois delegation is allocated in this manner.¹⁰

Direct votes for these delegates occur at the congressional-district level as follows. In 2016, Illinois had 18 congressional districts, and each district was allocated three delegates to the Republican nominating convention. Before the election, each presidential campaign nominates three candidates for delegate in each of the 18 congressional districts. The ballot instructs voters to vote for up to three delegate candidates, who need not be bound to support

the institutional environment permit the same inferences as here. For example, in some other states, the delegate vote totals only determine which particular delegates have the opportunity to represent a particular candidate and not the total number of delegates that a presidential candidate gets. We are not aware of similar natural experiments in other countries.

¹⁰About 20 percent of the delegation is independently allocated to the delegates determined in two other ways: delegates set aside for prominent party leaders and delegates bound to the winner of the statewide “beauty contest.” These delegates do not appear on the ballot.

the same presidential candidate. For example, Figure 1 provides part of the relevant delegate-candidate section of the ballot from McLean County in the 18th congressional district of Illinois in the 2016 election. While the Jeb Bush and Chris Christie campaigns nominated three delegates with names voters likely perceived as white, the Donald Trump campaign nominated two likely-white delegates and one, Sandra Yeh, who voters likely perceived as nonwhite—0 out of 30 Amazon Mechanical Turk workers we showed this name expected Yeh to be white. A Trump voter who pays “psychic costs” in voting for nonwhites could vote for only Trump’s two likely-white delegate candidates and not for Yeh. The performance of Yeh versus her white counterparts on the same ballot is thus one “natural experiment” of the 816 we observe, within which we analyze variation in delegate vote totals.

As Figure 1 shows, names of delegates are printed in large, bold font, followed by the last name of the presidential candidate to whom they are bound. Ballots provide no other information to voters. Delegates are grouped by presidential candidate, such that voters can easily identify delegates bound to their preferred presidential candidate.¹¹

We exploit three unique features of this setting to study discrimination. First, to identify discrimination, the presence of white delegate candidates running alongside the non-white delegate candidates who appear on the same ballots, in front of the same voters, bound to the same presidential candidates provides a naturally-occurring control group for estimating discrimination. Second, indicating tastes as the likely mechanism for any observed discrimination, there are essentially no incentives for rational voters to engage in statistical discrimination. Since delegates have no meaningful discretion if elected and are merely a mechanism for voting for a given presidential candidate, delegate names vary the “psychic cost” (Becker, 1957) of voting for a possibly-nonwhite delegate without varying relevant information about the consequences of that vote, which are plausibly fixed within the individual “experiments” we analyze. In other elections, voters may rationally infer nonwhite candidates differ on other dimensions, such as ideology or competence; here the only relevant dimension is the presidential candidate to which the delegate is bound, which is printed on the ballot. Third, any such discrimination entails trade-offs—the essence of taste-based discrimination—as it undermines voters’ preferred presidential candidates’ nomination prospects.¹² The election design therefore creates opportunities to credibly identify voter discrimination in this context—and, given the costs of discrimination and the natural absence of incentives for statistical discrimination, suggest taste-based discrimination as the

¹¹The same electoral process occurs for alternate delegates, for whom voters vote simultaneously. Our sample pools normal and alternate delegates.

¹²To the extent voters receive expressive utility for voting for their chosen presidential candidate (Pons and Tricaud, 2018; Spenkuch, 2018), engaging in racial discrimination also denies them this utility. See Appendix B for further discussion.

likely mechanism. To fix ideas, we present a formalization of the voter’s decision problem in these elections in Appendix B. Section 4.4 and Appendix J discuss further the plausibility of alternative interpretations of the discrimination we observe beyond tastes.

2.3 Elections in the Dataset

A contested Illinois Republican presidential primary took place on March 21, 2000; February 5, 2008; March 21, 2012; and March 15, 2016. Except in 2000, these primaries occurred relatively early in the primary season, before a “presumptive nominee” was established but nevertheless with a clear front-runner. The median district-level contest was decided by only 2,541 voters. For further context on the elections we study, see Appendix C.

2.4 Candidates for Presidential Convention Delegates

There were 2,386 unique delegate candidacies in Illinois across the four presidential primaries included in this study. We drop from the sample six who ran as uncommitted to a presidential candidate and 62 whose names cannot be coded by gender, as we discuss below.

If voters are more likely to vote for delegates whose names they recognize, and if white or male delegate candidates are especially likely to hold other political positions that would generate name recognition, we may uncover spurious relationships between delegate candidates’ race, ethnicity, or gender and their vote totals. To assess this possibility, we obtained information on delegates’ backgrounds that some voters might plausibly know. In particular, we instructed research assistants to complete meticulously detailed “background checks” on all the delegates in our sample, searching for whether they have ever held public office, a local Republican Party leadership position, or other non-political posts. We then divided offices into four categories: major office (e.g. U.S. Representative), state legislature (e.g. State Senator), minor office (e.g. county board member, town mayor), or no office. Appendix D details our data collection procedure.

Table 2 reports the number of unique delegates in each category. 40.7 percent of delegates in our sample have held public office or similar leadership positions, the vast majority of whom were minor officials, such as members of their local Republican party committee or members of town councils.¹³ White delegates are not significantly more likely to be officials than nonwhite delegates, nor is their distribution across levels of office notably different from nonwhite delegates. However, male delegates are somewhat more likely to be officials than

¹³Due to data limitations, these positions reflect the highest office we detected at any time for a delegate. See Appendix Table A1 for descriptive statistics on the number of unique delegates holding each office broken down by specific office.

female delegates, and they are distributed differently across office levels, though the statistical significance of this difference is marginal. Considering the distribution over specific offices in Appendix Table A1, however, we reject the null of independence with respect to delegate race and gender. These results suggest a need to check our estimates of discrimination for bias due to differential officeholding, as we do in Section 4.3.

Table 2: Officeholding by Race and Gender of Delegate

Level of Office	Count and Percentage of Column				Total
	By Race		By Gender		
	White	Nonwhite	Male	Female	
Major Office	18	4	19	3	22
	0.84	2.22	1.14	0.47	0.95
State Legislature	166	11	131	46	177
	7.76	6.11	7.83	7.14	7.64
Minor Office	687	57	553	191	744
	32.13	31.67	33.03	29.66	32.10
No Office	1,267	108	971	404	1,375
	59.26	60.00	58.00	62.73	59.32
Total	2,138	180	1,674	644	2,318

Notes: This table reports the distribution of the delegate population across four office categories by race and gender. Major offices include the governorship and U.S. House of Representatives membership; minor offices include City Council or Board of Education seats. We dichotomize the MTurk race measure and SSA gender measure at 0.5. The $\chi^2(4)$ statistics by race and gender are respectively 3.95 ($p = 0.27$) and 5.88 ($p = 0.12$). See Appendix Table A1 for a detailed tabulation of delegates by type of official and race: Results differ in the test of independence with respect to gender but not with respect to race. See Appendix D for a discussion of the data collection process.

2.5 The Illinois Republican Primary Electorate

Mirroring national patterns, Illinois Republican primary voters appear to be almost entirely non-Hispanic whites but are approximately balanced with respect to gender. Two separate data sources suggest this same conclusion. First, Appendix Table A2 presents demographic summary statistics on Republican primary voters in Illinois from complete-count administrative voter records from the Illinois Secretary of State with demographics estimated by the firm Catalist. Using their names and neighborhood racial composition, Catalist estimates that, of the people that official records indicate voted in the 2008, 2012, or 2016 Republican primaries, over 95 percent are non-Hispanic white and 51 percent are men. Data from the

2000 election are unavailable. Second, of Illinois Republicans who participated in a large-sample national survey, the Cooperative Congressional Election Study ([Ansolabehere, 2010](#)), and who administrative records show voted in the 2008 primary, 97 percent indicated they were white, and 55 percent indicated they were men ($N = 189$ Illinois Republicans who voted in the 2008 primary).

Given the paucity of nonwhite voters, these elections provide little opportunity to distinguish bias towards coethnics from bias against nonwhites by whites and nonwhites, but also largely eliminate the risk of attenuation of our estimates due to opposite-sign coethnic biases. Conversely, because of statewide and county-level approximate gender balance, we emphasize our results for gender discrimination are *net* results and that we are mute as to whether men and women discriminate in favor or against co-gender delegates.

Catalist data also indicate mean voter age was 60. Voters lived in Census block groups where, on average, one third of residents were college graduates and median annual per capita income was about \$70,000. These averages fit with national data which finds that primary voters tend to be whiter, older, more educated, and higher-income than nonvoters.

Propitiously for the external validity of our findings, available data suggest Illinois seems not far from the median U.S. state in terms of racially-discriminatory tastes. While the strength of racial tastes do not lend themselves to easy quantitative measurement, Appendix Table [A3](#) reports data on rates of racially-charged Google searches, 2.4 million results of self-administered Race Implicit Association Tests, the per-capita number of active hate organizations identified by the Southern Poverty Law Center, and the per-capita rate of race-related hate crimes as reported by the Federal Bureau of Investigation. None of these measures identify Illinois as a state with unusually high or low levels of racial animus.

Illinois Republican voters appear to place utility on voting for their preferred presidential candidates in these contests. We show this in two ways. First, comparing presidential “beauty contest” and delegate vote totals, Appendix Figure [A2](#) shows that delegates receive about 84 percent of the votes cast for their candidate in the “beauty contest” totals, suggesting the vast majority of “beauty contest” voters also participate in delegate contests. Second, Appendix Table [A4](#) shows that primary election turnout was 4–7 p.p. (18–27 percent) higher on average as a share of the voting-eligible population in presidential than in non-presidential election years from 1980 to 2016, controlling for Senate election years, even though the non-presidential years during this period featured contested Illinois governor primaries and the presidential years did not.

3 Data

3.1 Vote Totals

We observe official vote counts by delegate candidate at the county–congressional district level for every delegate candidate and county–district in Illinois in 2000, 2008, 2012, and 2016. A “county–congressional district” refers to the intersection of a county and congressional district: Some congressional districts span multiple counties and we observe vote totals separately within these county intersections for each delegate.¹⁴ These county–district intersections are mutually exclusive and exhaustive of Illinois voters and geography. Our unit of observation is each delegate and county–district. Importantly, we do not observe voting at the ballot level, and so we cannot study individuals’ joint voting decisions, nor decisions to “undervote” (i.e., not exhaust all N votes for delegates) versus spreading votes among the delegates of multiple presidential candidates.¹⁵

Our sample spans 2,318 unique delegate candidates and 19,711 vote-count observations, as we observe how a delegate candidate performed in multiple county-congressional district intersections, representing a total of 22.3 million votes. The mean (median) delegate candidate received 1,133 (306) votes in each county-congressional district area. All vote-count data were drawn and are publicly available from the online database of the Illinois State Board of Elections. The data also include the name of the delegate candidate as printed on the ballot, which is fixed at the congressional district level.

Throughout this paper, we refer to delegate candidates who run in the same county–district, in the same year, for the same presidential candidate, and for the same potential set of regular or alternate delegate positions as in the same “cell.” Recall that to maximize the value of their vote voters should vote for all of their preferred presidential candidates’ delegate candidates in the same cell and have no votes left after that. Delegates in the same cell are the most suited to comparison, in that factors related to geography, year, and presidential candidate are all constant within cells and that all remaining variation is between delegate candidates. In all our specifications we include fixed effects for each cell.

3.2 Inferring Delegate Race and Ethnicity

We measure how voters should perceive delegate candidates’ race and ethnicity from their names in three ways: using 2000 U.S. Census data, using a proprietary anthropological

¹⁴The same geographic unit is used, for example, in [Autor et al. \(2016\)](#). Total votes by district, not county, determine election outcomes. A delegate candidate can only run in one congressional district.

¹⁵Appendix Figure [A2](#) shows that delegates receive about 84 percent of the votes cast for their candidate in the “beauty contest.”

database of full-name frequencies (Onolytics), and using guesses of workers on Amazon Mechanical Turk (MTurk). These three sources yield two objective measures of likely race and ethnicity (Census and Onolytics) and one measure of subjective racial and ethnic perception (MTurk). For all three measures, we distinguish between delegates who are white, black, Hispanic/Latino, and Asian; for Onolytics and MTurk, we further make further distinctions among Asians. Throughout the paper we present estimates using all our measures of the racial and ethnic information in voters’ names to demonstrate the robustness of our results.

3.2.1 Census Data on Last Names

Public-use tabulations from the 2000 U.S. Census report, for each last name occurring 100 or more times in Census returns, the count and racial and ethnic composition of individuals with the last name. The tabulations include data for 151,671 last names and define racial categories as non-Hispanic white only, black only, Hispanic only, Asian or Pacific Islander only, and several other smaller categories. We use only the specified four categories.¹⁶

Similar to [Fryer and Levitt \(2004\)](#), we match delegates in our sample with the Census racial-composition data. Our measure is the racial composition of the U.S. population with the same last name and is thus continuous. About 87 percent (2,073 of 2,380) of delegates’ last names match an entry in the Census data exactly. For the remaining names, we identify the nearest match in the Census data for each delegate last name by minimizing the Jaro–Winkler distance, a common measure of string similarity in record matching. Appendix Table [A5](#) presents estimates including these inexact matches. The results remain similar, confirming that our results are unaffected by omitting delegates with rare names.¹⁷

3.2.2 Onolytics Classifier of Full Names

We also use a commercial software package to estimate the races of the delegate population. Onolytics is developed in [Mateos \(2014\)](#) and classifies names by a proprietary international database of over 1 million last names and 500,000 first names. While Onolytics provides detailed ethnicity categories, we collapse these to seven: black, Asian, Hispanic/Latino,

¹⁶There is significant missingness of racial-composition data for the smaller racial categories in the less-common last names in the tabulation and no delegate-candidate last names in our sample matched with last names that Census data identified as substantively (10 percent or more) linked to these smaller racial categories. These racial categories we omit are: American Indian or Alaska Native only, Native Hawaiian or Other Pacific Islander only, and Two or More Races.

¹⁷One advantage of the Census measure’s reliance on last names only is that the racial signals last names send are not contaminated by class signals individuals’ parents might have sent when choosing racially distinctive first names ([Fryer and Levitt, 2004](#)). This is also less of a problem in our context, as Asian and Hispanic names send opposite class signals, yet we find discrimination against both groups.

Indian, Middle Eastern, non-Hispanic white ethnic, and non-ethnic white.¹⁸ We define the nonwhite categories to correspond as closely as possible to those in our other two measures of delegate race and ethnicity. The measure is dichotomous.

3.2.3 MTurk Perceptions of Full Names

To measure voter-perceived race of delegates, we paid MTurk workers to guess the race of delegates from their provided full names. We followed the procedure of [Kuziemko et al. \(2015\)](#) to ensure high-quality guesses, in particular limiting the sample of potential participants to “master” MTurk workers who live in the United States. We paid for 30 guesses for each delegate name to yield reasonably precise estimates of perceived race. Our measure of race is the racial composition of these guesses and is thus continuous. See Appendix F for MTurk survey details and Appendix I for an analysis of attenuation bias due to measurement error.

An advantage of the MTurk measure is we could ask MTurk workers to provide their perceptions in finer ethnic categories than available from the U.S. Census. MTurk workers coded the ethnicity of each delegate as one of six categories: white, black, Hispanic, Asian, Indian, or Middle Eastern. By comparison, the Census definition of “Asian or Pacific Islander” spans individuals of East Asian, Indian, Middle Eastern, and Pacific Islander heritage.

3.3 Inferring Delegate Gender and Age

We also capture information about delegate gender and age contained in first names.

American first names robustly predict gender. To map delegates to likely genders, we use the baby-name file of the Social Security Administration (SSA) from 1930 to 2012, which covers all individuals born after 1930 and issued a Social Security card. 95 percent of delegates have first names that are either more than 95-percent male or more than 95-percent female in the SSA data. 62 delegate first names cannot be gender-coded and are dropped from our main analysis: These first names are either entirely missing from the ballot and thus our data or cannot be exact-matched in the SSA file. We use a continuous estimate of likely gender, the probability that a delegate is female is the proportion of U.S. citizens born with the same first name who are female at birth.

¹⁸See Appendix E for the details on the collapse. We define the white-ethnic category motivated by historical evidence for discrimination against “white ethnics” and for the attenuation of social distinctions among whites in the U.S. in the 20th century (e.g., [Roediger, 2005](#)). We define white ethnics in terms of the Onolytics classification for names of Southern European, Eastern European, and Jewish origin. The non-ethnic white category therefore includes names that are of Northern European, Central European, English, or Celtic origin. We generally collapse “white ethnic” and white names but also present them separately as a robustness check.

First names are also informative about age due to changes over time in the relative name popularity. To map delegates to their likely ages, we use the SSA baby-name file to find the year in which a name attained its maximal share of births. About half of all births occur within 10 years of this “modal year.” Among delegates, the median modal year is 1955, with a standard deviation of 20 years.¹⁹

3.4 Measuring Other Delegate Attributes

We also gathered information on delegates that voters can less readily infer from names and, as such, that they may be less likely to know. We determined delegates’ home counties and Census block groups from the residential addresses reported in official candidacy filings. We use these data in two ways below. First, they allow us to restrict the sample to voters who, due to physical distance, are least likely to have information about a delegate beyond what is reported on the ballot. Second, we use block-group characteristics as a proxy for delegate socioeconomic status: in particular, the share with at least a bachelor’s degree and log median household income.

3.5 Descriptive Statistics on Delegate Race and Ethnicity

To illustrate the racial and ethnic information in delegate names, Table 3 reports the five names identified as whitest and least white using the Census and MTurk whiteness measures; the dichotomous Onolytics measure does not enable such a ranking. Both measures identify names such as “Carol Hornickle” and “Mike Marron” as white and “Baba Padmanabhan” and “Noella Chung” as nonwhite. In Appendix Figure A5, we plot kernel density estimates for these measures. Most names in the sample are identified as very likely belonging to non-Hispanic whites, with a heavy left tail of names that likely belong to nonwhites.²⁰

To increase power over any individual race measure, we use an index constructed by estimating the first principal component (PC1) of the three race measures, rescaled to the unit interval to permit interpretation, as our baseline measure of delegate-candidate race throughout the results presented in Section 4. We report the results from the principal component analysis, including for the detailed race categories, in Appendix Table A7. Overall, we find that same-race, different-measure correlations—for instance, the MTurk black measure’s correlation with the Census black measure—are robustly positive. Our interpretation of these results is that all three measures capably differentiate between white and nonwhite names, and among detailed nonwhite categories, but with substantial noise that will bias

¹⁹See Appendix Figures A3 and A4 for histograms of SSA percent female and modal year of birth.

²⁰Appendix Table A6 gives examples of highly suggestive names for all racial and ethnic categories.

Table 3: Whitest and Least White Names of Delegate Candidates

	Census		MTurk	
	Name	Whiteness	Name	Whiteness
Whitest	Carol Hornickle	0.9946	Jill Bess	1
	Brian Milleville	0.9956	Helen Manson	1
	Sherry Hellmuth	0.9958	Mike Marron	1
	Ralph Baahlmann	0.9972	William S. Graham	1
	Gregory Musinski	0.9942	David L. Snyder	1
Least White	Baba Padmanabhan	0.0141	Noella Chung	0
	Ji Chung	0.0234	Angel Garcia	0
	Neil V. Patel	0.0155	Gustavo Gonzalez	0
	Noella Chung	0.0234	Rafael Rivadeneira	0
	Steve H. Kim	0.0260	Raja Sadiq	0

Notes: This table lists the five whitest and least white delegate names using the two continuous race measures in this paper. For the Census data, whiteness is defined as the proportion of U.S. citizens with the delegate’s last name who are non-Hispanic white. For the MTurk data, whiteness is defined as the proportion of Turkers who perceive the full name as non-Hispanic white. Ties are resolved by random draws. The categorical definition of Onolytics race variable means there is no equivalent ranking of names by informativeness. For further detail by race category, see Appendix Table [A6](#).

our estimates toward zero, motivating the use of principal component analysis to extract the common signal.

Using the modal guesses of delegate race from the MTurk data, the delegate population is 94 percent non-Hispanic white, 1 percent black, 4 percent Hispanic, and 1 percent Asian. Appendix Table [A8](#) reports further summary statistics on delegates. Appendix Table [A9](#) shows that likely-nonwhite delegates are nominated by nearly all the presidential candidates in our sample. Our results are thus not driven by a single presidential candidate’s voters.

3.6 Where and Why Minority Delegates Are Nominated

Campaigns recruit and nominate delegates. This introduces the potential for two varieties of selection. First, the nonwhite or female delegates that candidates select may differ from the white and male delegates. We take up this concern in Section [4.3](#). Our results weigh strongly against the presence of important confounding delegate-level unobservables.

Another form of selection is that campaigns may be more likely to nominate nonwhite or female delegates in areas of Illinois with specific characteristics. This presents a threat to external validity but, importantly, would not bias our results due to the presence of fixed effects which restrict our comparisons to within-cell variation. We evaluate this external-validity concern in Appendix [G](#). We regress the shares of nonwhite and female delegates on

several county–district observable characteristics, including the white and college-educated shares of population, white per-capita income, and the Republican two-party vote share. We find that campaigns frequently nominate nonwhites and women throughout the state, although relatively more frequently in less-Republican areas. We detect no other differences. The degree of selection of cells into the identifying set is sufficiently mild that the set of cells that contribute to identification closely matches of Illinois on average. In addition, we provide direct evidence that selection on county–district observables is not a threat to external validity with respect to Illinois statewide. Using coarsened exact matching, we reweight the sample so that cells with likely-nonwhite delegates match Illinois statewide on the four observables above. We estimate a similar penalty against nonwhites, implying that nonwhites are not more likely to run in areas where voter discrimination is lower.

If voters discriminate against nonwhite candidates, why would campaigns nominate them? In discussions with several officials responsible for recruiting delegate candidates, we found that recruitment costs were a common explanation, consistent with our finding in Appendix G that nonwhites were more likely to be nominated in less-Republican areas, where the supply of Republicans who could serve as delegate candidates may be more limited.²¹

4 Racial and Gender Discrimination by Voters

4.1 Empirical Strategy

To estimate the effects of discrimination on voting behavior, we compare the vote totals of delegates who differ in race, ethnicity, or gender but are in the same “cell”: delegates running in the same county–district, in the same year, for the same presidential candidate, and for the same potential set of regular or alternate delegate positions—all of whom voters should select to maximize the value of their ballot. Our baseline Poisson regression specification is:

$$\mathbb{E}[\text{Votes}_{ipct}] = \exp \left(\beta \cdot \text{Nonwhite}_i + \gamma \cdot \text{Female}_i + \mathbf{X}'_{ipct} \delta + \alpha_{pct} \right), \quad (1)$$

where Nonwhite_i and Female_i are our proxy variables for whether voters believe delegate i is nonwhite and female, respectively. α_{pct} is a vector of cell fixed effects, where p denotes the

²¹Campaigns must secure three supporters to run in every congressional district months before the primary takes place, and these delegates must agree to pay their own travel and lodging to attend the convention if elected. Campaigns therefore face search costs in locating delegates willing to serve. In addition, Illinois campaign managers as agents may not fully internalize the presidential candidates’ incentives. For example, the Trump campaign appears to have recruited from a campaign email list (Brueggeman, Brian. 5 March 2016. “Meet your delegates: the people who will vote for presidential candidates at the conventions.” *Belleville [IL] News-Democrat*). Appendix H further argues delegate service is a form of hobbyist consumption.

presidential candidate, c the county–district, and t the election year. For our estimates of β and γ to be unbiased, race and gender must be uncorrelated with unobservable determinants of votes. To show robustness, we add a vector of covariates \mathbf{X}_{ipct} in some specifications.

Estimates of β and γ can be interpreted as the average percentage of votes lost or gained due to discrimination by race and gender. Since all our race measures are scaled to the unit interval, the coefficients represent the estimated difference between a certainly-white and a certainly-nonwhite candidate. In alternative specifications, we replace Nonwhite_i with variables for specific nonwhite race and ethnicity groups. The dependent variable is the vote count for a delegate, and the unit of observation is the county–district–delegate–year.²² Standard errors are clustered at the delegate level, as this is the level at which the “treatment,” a delegate’s race, is assigned.²³ The presence of cell fixed effects in all regressions ensures that coefficients reported in the tables only exploit variation in the performance of delegate candidates within the same cell.

4.2 Main Results

Table 4 presents our main results. Column 1 reports estimates of Equation 1 using the share of MTurkers who perceived each delegate as nonwhite. The coefficient implies that if two delegates were running in the same cell but all MTurkers perceived one as white and the other as nonwhite, the latter would receive approximately 8.7 percent fewer votes in these elections. Column 2 breaks down these estimates by delegate race and ethnicity. Due to the small number of likely-black delegate candidates, our estimate of discrimination against black candidates is relatively imprecise. Delegates with names MTurkers perceive as Asian and especially as Indian and Middle Eastern appear to perform much worse than their counterparts running to represent the same candidates, receiving about 15 percent fewer votes than white delegates in the same cell. Using our SSA data on the female share of first names to code delegates who are objectively likely to be female, we find little evidence for discrimination against women in both Columns 1 and 2. Conditional on being in the

²²Other specifications, such as unweighted OLS on the number of votes or on log votes, would not estimate a meaningful quantity of interest due to heterogeneity in the number of votes by presidential candidate. See Appendix Tables A10 and A11 for OLS regressions with two alternative dependent variables, respectively vote shares and $\ln(1 + \text{votes})$, weighted by votes. Both find similar results.

²³We cluster standard errors at this level because the implied randomized experiment is that the same individual delegate candidates were randomly assigned to switch race or gender with other delegate candidates running in their same cell. We present a permutation test later in the paper that implements this implied experiment under the sharp null hypothesis. Appendix Table A14 reports our main results clustering at the level of congressional district, presidential candidate, alternate or regular, and year, which replicates the “slate” of delegates chosen by the same presidential candidate and competing for the same position. We find this increases standard errors by only about 20 percent on average and thus leaves the significance of our results almost entirely unchanged.

Table 4: Effect of Delegate Candidate Race and Gender on Votes, by Race and Ethnicity Measure

	MTurk		Census		Onolytics		Rescaled PC1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.006 (0.006)	0.006 (0.006)	0.004 (0.006)	0.001 (0.006)	0.005 (0.006)	0.004 (0.006)	0.005 (0.006)	0.001 (0.006)
Nonwhite	-0.087*** (0.016)		-0.045*** (0.014)		-0.035*** (0.007)		-0.092*** (0.016)	
Black		-0.033 (0.050)		0.018 (0.025)		-0.040*** (0.015)		-0.094*** (0.032)
Hispanic/Latino		-0.057*** (0.018)		-0.061*** (0.018)		-0.045*** (0.012)		-0.079*** (0.018)
Any Asian				-0.069*** (0.014)				-0.110*** (0.022)
East Asian		-0.093*** (0.025)				-0.055*** (0.008)		
Indian		-0.174*** (0.040)				-0.076*** (0.013)		
Middle Eastern		-0.160*** (0.031)				-0.088*** (0.033)		
White Ethnic						-0.016* (0.009)		
N	18,958	18,958	16,945	11,166	18,639	18,639	16,668	11,049
Pseudo- R^2	0.991	0.991	0.991	0.992	0.991	0.991	0.991	0.992

Notes: This table presents the results of estimating Equation 1, yielding estimates of the percentage vote penalties by nonwhite race or ethnicity and by gender. “Any Asian” uses the Census definition of Asian race, which spans our subsequent categories of East Asian, Indian, and Middle Eastern. Column 7 provides our preferred estimates throughout this paper. In all regressions the dependent variable is the vote count for the delegate. The unit of observation is the county–district–delegate–year. All regressions include cell-level FEs. Delegates are defined as in the same cell if they run in the same county–district, in the same year, for the same presidential candidate, and for the same potential set of regular or alternate delegate positions. Sample sizes change because not all delegate names can be classified using the Census data or Onolytics algorithm. Appendix Table A13 shows versions of Columns 7 and 8 that estimate the female and race/ethnicity coefficients in separate rather than combined regressions. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

same cell, likely-female delegates receive approximately the same average number of votes as likely-male delegates. Appendix Table A15 shows that interactions between delegate race and gender are insignificant.

In Columns 3 and 4, we present our estimates using Census data to code the racial information in delegates’ last names. The sample is limited to delegates whose last names match the Census data exactly. Similar to the MTurk measure, we find that nonwhite delegate candidates receive fewer votes. On average, a delegate who was objectively likely to be nonwhite would receive approximately 4.5 percent fewer votes than a delegate objectively likely to be white. We find significant discrimination against delegates likely to be Hispanic or Asian, but Census data struggle to identify delegates likely to be black by last name.²⁴ We again find a tight zero for discrimination against women.

In Columns 5 and 6, we present estimates using the dichotomous Onolytics race categories. As we found using the MTurk and Census race variables, we estimate that delegate candidates identified by the Onolytics algorithm as nonwhite receive fewer votes. Broken down by ethnicity, we find significant shares of voters do not vote for black, Hispanic/Latino, Asian, Indian, and Middle Eastern delegate candidates. We also break out a “white ethnic” category but find only a weakly significant difference in votes between “ethnic” and “non-ethnic” whites. The null result for women is unchanged.

Columns 7 and 8 present estimates using the rescaled PC1 measure. In Column 7, which provides the preferred estimates of this paper, the coefficient on nonwhite implies that delegate candidates who are generally identified as nonwhite across the three measures receive 9.2 percent fewer votes.²⁵ We also obtain a precise point estimate of zero discrimination against female delegates. In Column 8, we break this result down by the three racial categories common across our three measures. We find significant discrimination against delegates of all three racial categories. We also treat these estimates as our preferred results for the detailed categories throughout the rest of the paper.

The large change in the black coefficient when using the PC1 index is due to the substantial measurement error in our measures for this category and by a small number of outlier observations that influence our rescaling of the PC1 index. Given these challenges, our estimates for discrimination against black delegates should be interpreted with caution. More generally, differences in sample coverage and differences in how variables measure race and

²⁴Although the mean percentage black by last name in Census data is 9.8 percent, the 95th-percentile black last name is only 35 percent black.

²⁵To guard against bias in our standard error estimates, we also conduct a permutation test in which we repeatedly re-randomize the PC1 race measure at the delegate level within district-presidential candidate-years. Appendix Figure A6 plots 10,000 draws from a Monte Carlo simulation of the main regression specification with treatment status permuted in this way. Our standard errors appear unbiased relative to the bootstrap. The estimated effect remains significant.

ethnicity explain differences in estimated coefficients.²⁶

The PC1 index is missing when any of its constituent parts are missing, as the Onolytics or Census exact measures sometimes are. The gender measure is also missing for some delegates, as previously noted. Appendix Table A16 shows that imputing missing values for the PC1 index and for gender does not change the results. In Appendix I, we correct our estimates for the attenuation introduced by measurement error, defined either narrowly in the sampling of the MTurk race measure or more broadly in using the first principal component of the three race measures to proxy for true perceived race, a latent variable.²⁷ Appendix Table A17 shows that our main results are essentially unchanged if the race and gender variables are dichotomized.

4.3 Robustness Checks

Our research design absorbs into fixed effects all attributes that may affect voting behavior, such as attributes of presidential candidates, but do not vary among delegate candidates who appear on the same ballots, in front of the same voters, in the same geographic areas, and bound to the same presidential candidates. A potential confound must therefore cause some delegate candidates to receive more or fewer votes than same-race or same-gender delegate candidates in the same cell and be correlated with delegate-candidate race or gender.

4.3.1 Differences in Local Political Networks

One such possible confound is a racial or gender differential in local political networks and serves to illustrate the main empirical challenge to our results. If white delegate candidates are better known to voters than nonwhite delegate candidates, for instance, and voters are more likely to vote for delegate candidates whose names they recognize, then the coefficient on candidate race would capture the effect of the racial differential in local political networks

²⁶For example, the Census race measure is constructed exclusively from delegate last names, although for some groups (e.g., blacks), first names are a stronger signal of race. In addition, the MTurk measure captures some differences between how individuals perceive race and ethnicity from names that diverge from objective data. For instance, a majority of MTurk respondents perceived a delegate with the last name “Mercadante” as Hispanic, but the name, per Onolytics, is Italian in origin and is coded by the Census as white.

²⁷We find that attenuation due to sampling is trivial. Attenuation due to noisy proxies for perceived race, however, may be substantial. A correction using Cronbach’s (1951) α implies that being perceived as nonwhite may reduce the number of votes a delegate receives by 11 percent. Being perceived as black, most notably, may reduce the number of votes by 41 percent. Estimates for Hispanics/Latinos and Asians, whose names more clearly indicate race, rise by comparatively less. These estimates require the strong assumption that disagreement among the race measures is entirely classical measurement error. If the non-common components of these measures affect voting with the same sign as the common component, the reliability correction will overstate the true magnitude of discrimination. The estimates we report in Appendix I are therefore most reasonably viewed as upper bounds.

as well as the direct effect of delegate-candidate race. A parallel logic could apply for gender. We take three approaches to addressing this possibility.

First, since delegate candidates who campaigns repeatedly decide to list across multiple elections are plausibly more likely to be public officials or have other unobservable qualities that would increase vote totals, we introduce a fixed effect for candidates who run in more than one election year in our sample as a lightweight test of whether voters have information about delegate candidates other than the name-implied race and gender that they use to determine their votes. Columns 1 and 2 of Table 5, which use the PC1 race measure, show that our results for both nonwhites and women are essentially unchanged when we control for repeat candidates. While repeat delegate candidates receive significantly more votes, repeat delegate status is not strongly correlated with either delegate race or gender.²⁸

Second, Columns 3–6 of Table 5 control for voters’ delegate-level prior information using our delegate “background checks,” as described in Section 2.4, which exhaustively collected information on other offices or political roles held by delegates. In Columns 3 and 4 we include indicators for each of the 17 types of offices we recorded. We find significant returns to officeholding, in line with the literature on candidate name recognition (e.g., [Panagopoulos and Green, 2008](#)). Accounting for differential officeholding by race and gender modestly increases our estimate of discrimination against nonwhites but, in contrast, suggests that women receive 2 percent more votes than comparable men, a significant difference.²⁹ In Columns 7 and 8, we drop from the sample every delegate for whom we were able to find had held or previously ran for any office, no matter how minor. Among the delegates about which voters plausibly know nothing other than the information on the ballot—their names and the presidential candidate to which they are bound—nonwhite delegates still receive about 10 percent fewer votes than white delegates. Such female delegates receive about 2–3 percent more votes than comparable male delegates.

It remains possible that white or female delegates are more likely to be highly-connected individuals in ways our “background checks” could not capture but that would increase their vote totals. We provide further evidence against this possibility by exploiting the facts that some congressional districts in our sample span large areas, often hundreds of miles from end to end, and that we can observe outcomes by county within each congressional district. Highly-connected individuals should benefit from connections principally in their home counties, and indeed we find delegates receive 8 percent more votes in their home counties as recorded in official candidacy filings. In Columns 7 and 8 of Table 5, we show

²⁸Appendix Table A18 reports the coefficients on the covariate terms.

²⁹This point estimate matches the small advantage female candidates enjoy on average in survey-based experiments, per a recent meta-analysis ([Schwarz and Coppock, 2019](#)).

Table 5: Robustness Checks

	With Controls For:									
	Repeat Candidates		Officeholders				Non-Home Counties Only		Ballot Order	
	(1)	(2)	Control		Non-Officeholders Only		(7)	(8)	(9)	(10)
			(3)	(4)	(5)	(6)				
Female	0.006 (0.006)	0.002 (0.006)	0.020*** (0.005)	0.020*** (0.006)	0.031*** (0.006)	0.019*** (0.006)	-0.002 (0.007)	-0.002 (0.006)	0.010* (0.005)	0.009* (0.005)
Nonwhite	-0.088*** (0.017)		-0.099*** (0.015)		-0.099*** (0.019)		-0.096*** (0.022)		-0.101*** (0.014)	
Black		-0.095*** (0.033)		-0.070*** (0.026)		-0.056 (0.040)		-0.081** (0.035)		-0.073*** (0.027)
Hispanic/Latino		-0.074*** (0.017)		-0.080*** (0.015)		-0.059*** (0.017)		-0.078*** (0.024)		-0.087*** (0.008)
Asian		-0.100*** (0.023)		-0.095*** (0.020)		-0.173** (0.083)		-0.224*** (0.057)		-0.113*** (0.022)
N	16,668	11,049	16,668	11,049	8,091	5,068	14,422	9,536	16,668	11,049
Pseudo- R^2	0.991	0.992	0.992	0.992	0.995	0.994	0.990	0.990	0.991	0.992

Notes: This table reports the results of estimating Equation 1 with varying sets of controls or sample restrictions. In all regressions above, the dependent variable is the vote count for the delegate and the race measure used is the rescaled PC1 measure. The unit of observation is the county–district–delegate-year. For coefficients on controls, see Appendix Table A18. All regressions include cell-level FEs. Delegates are defined as in the same cell if they run in the same county–district, in the same year, for the same presidential candidate, and for the same potential set of regular or alternate delegate positions. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

our results are robust to dropping delegates’ home counties, suggesting our results are likely robust to any racial or gender differences in unobservable local political networks.

In summary, these tests of observable and unobservable differences among candidates suggest that our estimates of discrimination are unlikely to be confounded by differences in officeholding or voter information. We find that our estimates of racial discrimination remain unchanged with controls for possible sources of this variation and that these possible sources appear uncorrelated with delegate race and ethnicity.³⁰ On the other hand, controlling for officeholding leads us to find significant, if small, discrimination towards women, rather than the null result before controlling for officeholding, as female delegates are less likely than male delegates to hold offices with substantial electoral returns in the context of this primary. We interpret these results as ambiguous with respect to whether there is no discrimination against women or discrimination towards them.

4.3.2 Ballot Order Effects

Another potential confound stems from ballot order effects (Miller and Krosnick, 1998), in which delegate candidates may receive more or fewer votes as a causal result of their ordinal position on the ballot. In particular, if presidential campaigns place nonwhite delegate candidates into their lowest positions on ballots, then the coefficient estimate on having a nonwhite name would be inflated by the indirect effect of ballot order. However, Appendix Table A12 shows that, conditional on the number of ballot slots available (2, 3, or 4), delegate race and ballot order are uncorrelated, as are delegate gender and ballot order. We augment our specification in Equation 1 with controls for ballot order using dummy variables for the rank (1–4) of a delegate among those in the same cell and interact these with the maximum number of delegates (2–4) for whom a voter may vote in a given congressional district and year. Columns 9 and 10 of Table 5 report these results. Our findings change little when controlling for ballot-order effects.

4.3.3 Other Types of Discrimination

Voters may have preferences over other attributes of candidates, such as age, education, or income. In Section 3, we describe our data on these attributes: We use the age that voters might infer from a delegate’s first name and the block-group characteristics of the delegate’s official residential address as proxies of their education and income. To the extent that voters know the socioeconomic status of delegates, these measures allow us to detect whether there

³⁰This implies nothing about whether nonwhites are unconditionally more or less likely to serve in government or party offices in Illinois—only that, conditional on being nominated as a delegate, nonwhites and whites have similar officeholding profiles.

is discrimination on these dimensions. Appendix Table [A19](#) reports our results: We find precise nulls on discrimination by estimated age, education, and income.

These results suggest our estimates of racial discrimination are unlikely to be contaminated by discrimination on characteristics correlated with race, such as age or socioeconomic status. [Fryer and Levitt \(2004\)](#) demonstrate this threat to inference by showing that black Americans whose parents give them racially-distinctive first names are lower on average in socioeconomic status. In our context, this concern is relatively unlikely a priori. Our estimates are mostly driven by racial minorities other than blacks and remain robust when we use the Census measure, which is based on last names only, indicating that inferences about status from distinctively-black first names do not drive the results. Furthermore, our estimates for discrimination against East Asians and Indians—who have higher median income than whites on average both in the U.S. in general and in Illinois specifically—are still significant and negative.

4.4 Residual Incentives for Statistical Discrimination

While the design of the primary suggests taste-based discrimination as the likely mechanism for our findings, we discuss here the possibility that voters have residual incentives for statistical discrimination. In Appendix [J](#), we also discuss other alternative interpretations: (1) that voters are making inferences about presidential candidates from delegates, (2) unobserved confounds in general, (3) the implications of voter indifference between Presidential candidates, (4) whether voters use delegate voting to send signals to presidential candidates or party elites, and (5) whether voters understand whether delegate voting has stakes. There we also report an original survey of self-identified Illinois Republican primary voters.

An interpretation of the discrimination we observed as taste-based relies upon the assumption that rational voters have minimal incentives to engage in statistical discrimination. Following [Becker’s \(1957\)](#) definitions of taste-based and statistical discrimination as with respect to how one would interpret observed behavior as if it were undertaken by rational agents, it is difficult to see why a rational agent in this setting would perceive incentives for statistical discrimination. Even a rational agent who misunderstood the institution, not knowing that delegates were bound, would need to maintain very unlikely beliefs: A substantial fraction of rational agents would need to believe that white delegates selected by the opposing candidates would be more likely to vote for their candidate of choice at the convention than nonwhite delegates selected by their candidate of choice. Moreover, as the elections we study had largely narrowed to two contenders, voters who believed delegates were likely to abstain or defect to their preferred presidential candidate’s rival have only

the alternative of helping that rival. Under such conditions, it would not be sufficient for a voter to believe nonwhite delegates were more likely to defect than white delegates for the same presidential candidate. For example, to explain our results in 2012, rational voters would need to believe both that delegates were able to change conventional rules and exert discretion and that, once doing so, the nonwhite delegates Mitt Romney had selected were less likely to vote for Romney at the convention than the white delegates Rick Santorum had selected were to desert Santorum in order to vote for Romney.³¹ That convention rules do not allow delegates to engage in this behavior in the first place makes it all the more implausible.³²

Several pieces of empirical evidence also are inconsistent with the presence of statistical discrimination. First, that our estimates do not meaningfully vary with the ideological positions of the candidates—as measured by ideology scores compiled by [Bonica \(2013\)](#), as we show in [Section 5](#)—is inconsistent with voters inferring nonwhites are more likely to abandon conservative presidential candidates to vote for liberals; voters voting for more liberal Republican presidential candidates discriminate no less. This result also suggests our findings are unlikely to come from voters concerned that nonwhite delegates would move party ideology left or right through their presence at the convention or in potential future political careers. Second, in [Appendix J](#) we present results from a survey we conducted of a convenience sample of Illinois Republican primary voters that finds that perceived differences in the loyalty of white versus nonwhite delegates are much too small to plausibly explain our finding of discrimination against nonwhite delegates. In contrast, about 10 percent of voters instead said they would avoid voting for nonwhite delegates because they were “uncomfortable” doing so, despite potential social desirability bias, consistent with “psychic costs.” Third, although the risk of further rounds of convention balloting (sometimes called a “contested convention”) was heightened in 2016, [Table A25](#) finds that voter discrimination does not vary meaningfully across years, including in 2000, 2008, and 2012 when the risk of further rounds of balloting at the convention was not considered plausible by the time

³¹A corollary of this observation is that even if one were to adopt a broader definition of statistical discrimination that includes choices stemming from mistaken beliefs ([Bohren et al., 2019](#)), it seems unlikely that behavioral voters with such beliefs would perceive incentives to discriminate.

³²The sole case in which delegates do have discretion is if the convention is contested and goes to a second round of voting. For example, in the 2016 presidential nominating contest, some observers anticipated a possibility that no presidential candidate would receive a majority of votes on the first ballot at the convention, “un-binding” delegates for subsequent rounds of voting. However, as reviewed in [Appendix C](#), the primary race had progressed sufficiently to a two-candidate race by the time Illinois voted in 2000, 2008, and 2012 that multiple rounds of balloting were essentially impossible, and our estimates remain largely unchanged when examining the two frontrunners in these years only. Moreover, the last contested Republican convention was in 1952, and the last one close to contestation was in 1976. Given the elections in our dataset, it seems a priori unlikely that voting behavior in Illinois was informed by the possibility of second-round convention balloting.

Illinois voted. Finally, our discussion of the nature of convention in Appendix H shows that voters who discriminate are unlikely to be trying to prevent political power from accruing to nonwhites.

5 Heterogeneity in Discrimination

We next estimate how racial and gender discrimination varies along several dimensions. Although some of these tests have low statistical power and we caution that they are fundamentally observational in nature, their results are consistent with a taste-based interpretation.

5.1 By Presidential Candidate Race and Gender

Would the discrimination we observe in these delegate elections also manifest in other voting decisions, such as in choices between presidential candidates themselves? If our estimates reflect discriminatory tastes, and if these discriminatory tastes also influence voters' choices of presidential candidates, we would expect presidential primary candidates who are nonwhite or female to attract voter populations with weaker tastes against nonwhites and stronger tastes towards women on average than the voter populations of white or male presidential candidates. We can test this prediction because the design of the primary allows us to separately estimate the magnitudes of race and gender discrimination among voters for each presidential candidate.³³

In Column 1 of Table 6, we show that voters for nonwhite presidential candidates indeed do not appear to have any racially discriminatory tastes on average, whereas voters for white presidential candidates do. Column 1 estimates Equation 1 with an interaction term for the race and ethnicity of the presidential candidate with the race of the delegate candidate as well as controls for delegate officeholding and ballot order. As with delegate candidates, for concision we define presidential candidates as white if they are non-Hispanic whites alone. We estimate that nonwhite delegates of white presidential candidates lost 10 percent of votes due to discrimination. A χ^2 -test rejects equality between delegates of white and nonwhite presidential candidates. This suggests that voters for nonwhite presidential candidates have, on average, weaker racial tastes than voters for white presidential candidates. Although we cannot rule out other explanations for this pattern, it is consistent with our estimates of discrimination as reflecting tastes that also affect voters' presidential candidate choices.

A similar result applies to women, consistent with the prediction that voters for female

³³Appendix Figure A7 presents the estimated level of discrimination among voters for individual presidential candidates.

Table 6: Heterogeneous Treatment Effects by Presidential Candidate Race and Gender

	(1)	(2)	(3)
Nonwhite Delegate			
× White Pres. Cand.	-0.101*** (0.017)		-0.104*** (0.017)
× Nonwhite Pres. Cand.	0.052 (0.064)		0.053 (0.069)
Female Delegate			
× Female Pres. Cand.		0.335*** (0.096)	0.335*** (0.109)
× Male Pres. Cand.		0.003 (0.006)	0.003 (0.007)
χ^2 test of equality	7.07***	12.49***	15.73***
N	17,126	18,958	16,668
Pseudo- R^2	0.990	0.991	0.991

Notes: This table reports the results of estimating Equation 1, interacting with the race and gender of presidential candidates. We code Ben Carson, Ted Cruz, Alan Keyes, and Marco Rubio as nonwhite (including Hispanic), Carly Fiorina as female, and the remaining 19 candidates as white men. The χ^2 test results also include interactions of the nonwhite and gender variables with the year, thus exploiting only within-year comparisons of presidential candidates by attribute. In all regressions the dependent variable is the vote count for the delegate. The unit of observation is the county–district–delegate–year. All regressions include controls for ballot order and detailed officeholding and cell-level FEs. Delegates are defined as in the same cell if they run in the same county–district, in the same year, for the same presidential candidate, and for the same potential set of regular or alternate delegate positions. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

presidential candidates may similarly be selected towards relatively pro-female tastes. However, there was only one female Republican presidential candidate whose delegates won any votes in Illinois from 2000 to 2016, Carly Fiorina in 2016, and so our results come with the immediate caveat that they rely upon vote totals for one female candidate. Comparing the vote totals of female and male Fiorina delegates in the same cell, we find substantial discrimination towards female delegates: We show in Column 2 of 6 that female Fiorina delegates win about 35 percent more votes than otherwise-similar male Fiorina delegates, a statistically significant but imprecise estimate. The difference with non-Fiorina delegates is also significant in a χ^2 -test.³⁴

We also investigate whether discrimination varies by presidential candidate ideology, categorizing moderate and conservative presidential candidates on either side of the median

³⁴See Appendix Table A15 for evidence that the presence of Fiorina voters with the highest estimated bias against nonwhites and for women is not explained by missing interaction terms.

of ideology estimates inferred from the identities of campaign donors from Bonica (2013). Reported in Appendix Table A21, the relationship is insignificant, suggesting little sorting of voters who discriminate more according to left–right ideology. Importantly, this null result is also inconsistent with statistical discrimination wherein voters believe white delegates are more conservative than nonwhite delegates. Voters for moderate presidential candidates would have a weaker incentive to discriminate than voters for conservative presidential candidates, but we do not find this.³⁵

5.2 By Competitiveness and the Costs of Discrimination

Taste-based theories of discrimination predict individuals discriminate less when it is more costly for them to do so (Becker, 1957; Hedegaard and Tyran, 2018). In this setting, voters face a trade-off between any psychic costs of voting for nonwhites and the costs of engaging in discrimination—one instrumental component of which is the possibility that withholding their vote for their preferred presidential candidates’ minority delegates would prove decisive, causing the delegates to lose and impairing their preferred presidential candidates’ nomination prospects. Other components of the cost of discrimination, such as the intrinsic utility voters gain from the act of voting for their chosen candidate (Pons and Tricaud, 2018; Spenkuch, 2018), do not vary with the probability voters will be decisive here.

To evaluate this prediction, we split presidential candidates by whether they received above- or below-median shares of statewide votes in each respective election year. In our context, below-median candidates have essentially no chance of winning delegates, and thus their voters face lower instrumental costs of discrimination than voters for above-median candidates. Appendix Table A21 presents results consistent with a downward-sloping demand curve for taste-based discrimination: We estimate a 9-percent penalty for the nonwhite delegates of above-median presidential candidates, compared to a penalty of 57 percent for below-median candidates. A χ^2 -test confirms this difference is significant. When rational voters with tastes against nonwhites vote in an election they expect to be close, they appear to be more likely to prioritize the victory of their preferred presidential candidate over avoiding the “psychic cost” of voting for nonwhites than when no delegates are plausibly at stake. We interpret these patterns as suggestive evidence in favor of viewing the discrimination we detect as taste-based and as politically consequential even in competitive elections.³⁶

³⁵A lack of other proxies for ideology which cover Republican presidential candidates constrains further analysis of the relationship between ideology and discrimination.

³⁶Discrimination persists even among the top two presidential candidates’ voters in each year ($\beta = -0.087$, $p < 0.001$) as well as for the top presidential candidates’ voters ($\beta = -0.083$, $p < 0.001$). We repeat this and the ideology analyses for gender in Appendix Table A21. Preventing further analysis of voter responses to instrumental incentives, nearly all variation in pivot probability is at the candidate level, and

5.3 By Geography

We also examine whether our area-level estimates of racial discrimination vary with population attributes previous literature has identified: the share of the population that is white, the share of the white population that has a college degree, and the real median per capita income of whites. These are all measured by the U.S. Census at the county–district level. To estimate how they correlate with discrimination, we modify Equation 1 by interacting the nonwhite-delegate variable with our county–district-level demographic measures. In an alternate specification, we also add delegate-level fixed effects; since we observe vote counts across counties within delegates, adding delegate fixed effects allows electorate demography to vary while holding delegate identities constant.

Appendix Table A23 reports these results. There is no significant association between our county-level estimates of racial discrimination and the white share of population. Discrimination is lower in areas with higher college-educated white shares of population, with higher per-capita income of whites, and with lower two-party Republican presidential vote shares, although significance is often sensitive to specification. These findings are consistent with historical patterns wherein American political parties that explicitly appeal to racial prejudice tend to perform better in lower-income areas (e.g., Mulkern, 1990). We find little evidence that the effect varies with the Stephens-Davidowitz (2014) or Xu et al. (2014) measures of anti-black bias, perhaps because most of the nonwhite delegates in our sample are not black.³⁷

A parallel analysis of gender discrimination reaches similar conclusions. Appendix Table A24 finds that discrimination towards female delegates is stronger in areas where the college share of adults is higher and per-capita income is higher. These results, however, are not robust in within-delegate analyses. In lieu of local measures of gender bias, we consider three variables drawn from U.S. Census data: the adult sex ratio, the log male-female difference in annual labor earnings, and the male-female difference in the percentage of individuals over age 25 with at least a bachelor’s degree. We also consider the average Republican vote share.

within-candidate geographic variation is trivial. Another source of variation in the instrumental cost of discrimination is whether or not the delegate position is for a regular or an alternate delegate. See Appendix Table A22. We find only weak evidence for the hypothesis that racial and gender discrimination is greater against alternate delegates than against regular delegates. In tension with our results for competitiveness, the similarity of our estimates for alternates and delegates suggests an important role for disutility from the mere act of voting for a nonwhite delegate, irrespective of their probability of serving. Given these null results and the possibility that voters for presidential candidates who receive relatively few votes may have relatively stronger tastes against minorities, we regard this analysis of instrumental costs as suggestive.

³⁷The lack of correlation across these measures may be attributed variously to multidimensionality in the concept of racial bias, to distinctions between anti-black bias and bias against the other nonwhite groups examined in this paper, to measurement error in all three variables, or to potentially other factors. The correlation between the Stephens-Davidowitz (2014) and Xu et al. (2014) measures is only 0.1.

We find no association with gender discrimination for any of these variables in either the pooled or within-delegate analyses.

5.4 Over Time

We estimate Equation 1 but interacting the nonwhite and female variables with year dummies. Appendix Table A25 presents the estimates by year. An χ^2 -test of the equality restriction across years shows that we cannot reject either null hypothesis that discrimination is constant over time. We also impose linear time trends on racial and gender discrimination, starting from Equation 1, by interacting the variables with the continuous year variable. The point estimates on these interaction terms are essentially zero for both race and gender.

6 Discussion

Our estimate of racial discrimination in these elections is large, comparable in magnitude to the benefits in these elections of being a state legislator or running in one’s home county. This section reviews the political and policy implications of our findings.³⁸

We first evaluate the consequences of racial discrimination in Illinois Republican delegate selection by simulating counterfactual outcomes absent discrimination. To simplify, we assume delegates more likely than not to be nonwhite according to the MTurk race measure all lost the same fraction of votes due to discrimination and then vary the estimated penalty from 0 to -0.3 . We augment their vote totals by these fractions, which assumes that discriminating voters either undervote or proportionally vote for other delegates when nonwhites are nominated. We then calculate which delegates would have won under these augmented vote totals. This exercise suggests that discrimination reduced the number of nonwhites who served: Appendix Figure A8 shows that, at our preferred estimate of discrimination against nonwhite delegates of 9.2 percent, 5 additional nonwhite delegates would have won, relative to the baseline of 34. In addition, discrimination reduces the appeal of election winners on non-racial dimensions to voters who discriminate: We estimate that, due to discrimination, several nonwhite delegates lost to white delegates bound to other presidential candidates. Appendix Table A27 provides an example of one such likely change in outcomes. Voters in

³⁸We caution these are partial-equilibrium analyses. For instance, incentives of presidential candidates to nominate a nonwhite as a delegate, and those of nonwhite candidates to run for other offices, may also change in a counterfactual without voter racial discrimination. Plausible general-equilibrium effects in other elections, such as on the margins of nonwhite candidate entry or institutional design (Trebbe et al., 2008), could mean our estimate of the increase in nonwhite representation from the elimination of racial discrimination could be over- or understated.

this setting appear willing to discriminate despite the potential for discrimination to change outcomes and advantage disfavored presidential candidates.

Beyond the narrow context of Illinois Republican primaries, discrimination of the magnitude we observed would also be large enough to represent a substantial barrier to the election of racial minorities in other elections. Appendix Table [A26](#) presents back-of-the-envelope calculations that use U.S. House primaries from 1990 to 2010 to illustrate the potential magnitude of that barrier. As these calculations require the strong assumption that our estimate of discrimination in Illinois delegate-primary elections is valid for U.S. House primaries, we offer them to contextualize the substantive importance of our estimated magnitude of discrimination, not to reach precise conclusions about nonwhite representation in the U.S. House in a counterfactual world without racial discrimination. We estimate nonwhites would have won 19 additional Republican U.S. House primaries from 1990 to 2010 absent racial discrimination of the size we observed in this setting. This would result in an increase of about 9 percent in the number of nonwhites winning Republican primaries ($19/218 = 0.087$). These comparisons suggest racial discrimination is plausibly a critical barrier to minority political officeholding, which other research shows lies at the root of important political and economic racial disparities.

Would such discrimination manifest in other elections? Encouragingly for the external validity of our findings, we found that voters appear to select out of voting for nonwhite and female presidential candidates in a manner correlated with our estimates of their collective racial and gender tastes, consistent with discriminatory tastes having stakes for presidential candidate choices. Of course, many considerations influence whether discrimination is likely to be greater or smaller in magnitude in other elections. Both costs and benefits of discrimination to voters may vary. On the one hand, discrimination could be greater in other elections where voters may have less information or weaker preferences. For example, voters may know less about candidates in state legislative primaries than in presidential primaries and therefore have weaker candidate preferences; all else equal, we would expect greater discrimination there. Likewise, our estimates would not capture behavior arising from the anticipated psychic costs of having nonwhite officeholders, nor from any statistical discrimination against them. On the other hand, discrimination could be smaller in general elections where partisan preferences may be more important, or in elections where voters have on average weaker racial tastes. For example, in a general election between a nonwhite Republican and a white Democrat, Republican voters with strong racial tastes may still prefer paying the psychic costs of voting for the nonwhite candidate to voting for a Democrat.

In the U.S., credible evidence about discrimination in elections also has significant policy implications. The U.S. Congress and the Supreme Court regularly consult academic assess-

ments of voter discrimination in crafting and reviewing American election laws. Our results are most relevant to the debate over the Voting Rights Act of 1965, passed in part to facilitate the election of nonwhites to political office.³⁹ In the subsequent decades, Congress and the Court have evaluated the continuing necessity of the Act in part by attempting to answer whether nonwhite candidates still encounter discrimination in elections. In decisions in 2009 and 2013, the Supreme Court struck down portions of the Act, finding only “decades-old data” in support of federal claims of continued discrimination. Our findings are particularly applicable to vote-dilution cases under the Act in which plaintiffs seek the construction of majority–minority districts. First, our finding of racial and ethnic discrimination in voting behavior may meet evidentiary standards for injury often difficult to meet with existing correlational evidence. Second, by estimating significant discrimination against several nonwhite minority groups, we provide evidence of shared injury, often the relevant legal burden in “coalitional” cases brought jointly by minority groups under the Act.

Looking beyond the Act, our findings suggest policy responses that operate only on informational margins to reduce statistical discrimination, however effective (Casey, 2015), are likely to leave intact substantial barriers to the election of racial and ethnic minorities arising from taste-based discrimination. However, our work lends support to other work that finds political parties may be a more relevant constraint on female representation than voters (Baltrunaite et al., 2014, 2019; Esteve-Volart and Bagues, 2012; Casas-Arce and Saiz, 2015; Besley et al., 2017; Bhalotra et al., 2017; Schwarz and Coppock, 2019). The contrasting evidence that voters discriminate by race and ethnicity suggests that different interventions may be necessary to address nonwhite and female underrepresentation in political office.

References

- Akerlof, George, “The Economics of Caste and of the Rat Race and Other Woeful Tales,” *Quarterly Journal of Economics*, 1976, pp. 599–617.
- Ansolabehere, Stephen, “CCES, Common Content, 2008,” 2010.
- and Brian F. Schaffner, “CCES Common Content, 2016,” 2017.
- Anzia, Sarah F and Christopher R Berry, “The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?,” *American Journal of Political Science*, 2011, 55 (3), 478–493.

³⁹We thank Christopher Elmendorf and David Schleicher for helpful discussions about implications of our results with respect to the Act.

- Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi**, “Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure,” Working Paper, NBER 2016.
- Baltrunaite, Audinga, Alessandra Casarico, Paola Profeta, and Giulia Savio**, “Let the Voters Choose Women,” *Journal of Public Economics*, 2019, *180*, 104085.
- , **Piera Bello, Alessandra Casarico, and Paola Profeta**, “Gender Quotas and the Quality of Politicians,” *Journal of Public Economics*, 2014, *118*, 62–74.
- Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova**, “Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India,” *Science*, 2012, *335* (6068), 582–586.
- Becker, Gary S**, *The Economics of Discrimination*, Chicago: University of Chicago Press, 1957.
- Bertrand, Marianne and Esther Duflo**, “Field Experiments on Discrimination,” in “Handbook of Economic Field Experiments” 2017, pp. 309–393.
- and **Sendhil Mullainathan**, “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination,” *American Economic Review*, 2004, *94* (4), 991–1013.
- Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne**, “Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden,” *American Economic Review*, 2017, *107* (8), 2204–42.
- Bhalotra, Sonia, Irma Clots-Figueras, and Lakshmi Iyer**, “Pathbreakers? Women’s Electoral Success and Future Political Participation,” *Economic Journal*, 2017.
- Bohren, J Aislinn, Kareem Haggag, Alex Imas, and Devin G Pope**, “Inaccurate Statistical Discrimination,” Working Paper, NBER 2019.
- Bonica, Adam**, “Database on Ideology, Money in Politics, and Elections,” Stanford, CA: Stanford University Libraries 2013.
- Casas-Arce, Pablo and Albert Saiz**, “Women and Power: Unpopular, Unwilling, or Held Back?,” *Journal of Political Economy*, 2015, *123* (3), 641–669.
- Casey, Katherine**, “Crossing Party Lines: The Effects of Information on Redistributive Politics,” *American Economic Review*, 2015, *105* (8), 2410–2448.

- Center for American Women and Politics**, “Women in the US Congress 2016,” 2016.
- Chattopadhyay, Raghavendra and Esther Dufo**, “Women as Policy Makers: Evidence from a Randomized Policy Experiment in India,” *Econometrica*, 2004, 72 (5), 1409–1443.
- Clots-Figueras, Irma**, “Are Female Leaders Good for Education? Evidence from India,” *American Economic Journal: Applied Economics*, 2012, 4 (1), 212–244.
- Cronbach, Lee J**, “Coefficient Alpha and the Internal Structure of Tests,” *Psychometrika*, 1951, 16 (3), 297–334.
- Doleac, Jennifer L and Luke CD Stein**, “The Visible Hand: Race and Online Market Outcomes,” *Economic Journal*, 2013, 123 (572), F469–F492.
- Esteve-Volart, Berta and Manuel Bagues**, “Are Women Pawns in the Political Game? Evidence from Elections to the Spanish Senate,” *Journal of Public Economics*, 2012, 96 (3), 387–399.
- Ferreira, Fernando and Joseph Gyourko**, “Does Gender Matter for Political Leadership? The Case of US Mayors,” *Journal of Public Economics*, 2014, 112, 24–39.
- Fraga, Bernard L**, “Race, Party, and Candidate Prospects Across the Multiple Stages of Congressional Elections,” 2013.
- Fryer, Roland G and Steven D Levitt**, “The Causes and Consequences of Distinctively Black Names,” *Quarterly Journal of Economics*, 2004, pp. 767–805.
- Fujiwara, Thomas**, “Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil,” *Econometrica*, 2015, 83 (2), 423–464.
- Glover, Dylan, Amanda Pallais, and William Pariente**, “Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores,” *Quarterly Journal of Economics*, 2017, 132 (3), 1219–1260.
- Goldin, Claudia and Cecilia Rouse**, “Orchestrating Impartiality: The Impact of ‘Blind’ Auditions on Female Musicians,” *American Economic Review*, 2000, 90 (4), 715–741.
- Green, Donald, Bradley Palmquist, and Eric Schickler**, *Partisan Hearts & Minds: Political Parties and the Social Identities of Voters*, Yale University Press, 2002.
- Guryan, Jonathan and Kerwin Kofi Charles**, “Taste-Based or Statistical Discrimination: The Economics of Discrimination Returns to Its Roots,” *Economic Journal*, 2013, 123 (572), F417–F432.

- Hedegaard, Morten and Jean-Robert Tyran**, “The Price of Prejudice,” *American Economic Journal: Applied Economics*, 2018, 10 (1).
- Iyer, Lakshmi, Anandi Mani, Prachi Mishra, and Petia Topalova**, “The Power of Political Voice: Women’s Political Representation and Crime in India,” *American Economic Journal: Applied Economics*, 2012, 4 (4), 165–193.
- Kuziemko, Ilyana and Ebonya Washington**, “Why Did the Democrats Lose the South? Bringing New Data to an Old Debate,” *American Economic Review*, 2018, 108 (10), 2830–67.
- , **Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva**, “How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments,” *American Economic Review*, 2015, 105 (4), 1478–1508.
- List, John A.**, “The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field,” *Quarterly Journal of Economics*, 2004, pp. 49–89.
- Manning, Jennifer E.**, “Membership of the 114th Congress: A Profile,” CRS Report R43869, Congressional Research Service 2016.
- Mateos, Pablo**, *Names, Ethnicity and Populations*, Springer, 2014.
- Miller, Joanne M and Jon A Krosnick**, “The Impact of Candidate Name Order on Election Outcomes,” *Public Opinion Quarterly*, 1998, pp. 291–330.
- Mulkern, John**, *The Know-Nothing Party In Massachusetts: The Rise and Fall of a People’s Party*, Northeastern University Press, 1990.
- Pager, Devah, Bruce Western, and Bart Bonikowski**, “Discrimination in a Low-Wage Labor Market: A Field Experiment,” *American Sociological Review*, 2009, 74 (5), 777–799.
- Panagopoulos, Costas and Donald P. Green**, “Field Experiments Testing the Impact of Radio Advertisements on Electoral Competition,” *American Journal of Political Science*, 2008, 52 (1), 156–168.
- Pande, Rohini**, “Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India,” *American Economic Review*, 2003, 93 (4), 1132–1151.
- Phelps, Edmund S.**, “The Statistical Theory of Racism and Sexism,” *American Economic Review*, 1972, 62 (4), 659–661.

- Pons, Vincent and Clémence Tricaud**, “Expressive Voting and Its Cost: Evidence from Runoffs with Two or Three Candidates,” *Econometrica*, 2018, 86 (5).
- Price, Joseph and Justin Wolfers**, “Racial Discrimination Among NBA Referees,” *Quarterly Journal of Economics*, 2010, 125 (4), 1859–1887.
- Roediger, David R.**, *Working Toward Whiteness: How America’s Immigrants Became White*, Basic Books, 2005.
- Rubinstein, Yona and Dror Brenner**, “Pride and Prejudice: Using Ethnic-Sounding Names and Inter-Ethnic Marriages to Identify Labour Market Discrimination,” *Review of Economic Studies*, 2014, 81 (1), 389–425.
- Schwarz, Susanne and Alexander Coppock**, “What Have We Learned About Gender From Candidate Choice Experiments? A Meta-Analysis of 42 Factorial Survey Experiments,” 2019.
- Shafer, Byron E.**, *Bifurcated Politics: Evolution and Reform in the National Party Convention*, Harvard University Press, 1988.
- Spenkuch, Jörg L.**, “Expressive vs. Strategic Voters: An Empirical Assessment,” *Journal of Public Economics*, 2018, 165, 73–81.
- Stephens-Davidowitz, Seth**, “The Cost of Racial Animus on a Black Candidate: Evidence Using Google Search Data,” *Journal of Public Economics*, 2014, 118, 26–40.
- Teele, Dawn Langan, Joshua Kalla, and Frances Rosenbluth**, “The Ties That Double Bind: Social Roles and Women’s Underrepresentation in Politics,” *American Political Science Review*, 2018, 112 (3), 525–541.
- Trebbi, Francesco, Philippe Aghion, and Alberto Alesina**, “Electoral Rules and Minority Representation in US Cities,” *Quarterly Journal of Economics*, 2008, 123 (1), 325–357.
- Vogl, Tom S.**, “Race and the Politics of Close Elections,” *Journal of Public Economics*, 2014, 109, 101–113.
- Washington, Ebonya**, “How Black Candidates Affect Voter Turnout,” *Quarterly Journal of Economics*, 2006, pp. 973–998.

Xu, Kaiyuan, Brian Nosek, and Anthony Greenwald, “Psychology Data from the Race Implicit Association Test on the Project Implicit Demo Website,” *Journal of Open Psychology Data*, 2014, 2 (1).